COUNTING HYPERGRAPH COLOURINGS IN THE LOCAL LEMMA REGIME

Heng Guo (University of Edinburgh)
Joint with Chao Liao (SJTU), Pinyan Lu (SHUFE), and Chihao Zhang (CUHK)

Chinese Academy of Science, Dec 26 2017
COLOURINGS
Graph (proper) colouring

3-colouring of the Petersen graph
Phase transitions: as some parameter changes, macroscopic behaviours of the whole system change drastically.

E.g. ice → water → water vapor
Computational Phase Transitions

The complexity of determining whether a graph is $q$-colourable:

- $q = 1, 2$ : easy;
- $q \geq 3$ : \textbf{NP}-hard.

What about graphs with maximum degree $\Delta$?

- $q \geq \Delta + 1$ : always colourable;
- $q \geq \Delta - \sqrt{\Delta} + 3$ : polynomial-time (Molloy, Reed ’01 ’14);
- $q \leq \Delta - \sqrt{\Delta} + c$ : \textbf{NP}-hard (Embden-Weinert, Hougardy, and Kreuter ’98).

(c = 1.5 + $O(\Delta^{-0.5})$)
The complexity of determining whether a graph is $q$-colourable:

- $q = 1, 2$: easy;
- $q \geq 3$: NP-hard.

What about graphs with maximum degree $\Delta$?

- $q \geq \Delta + 1$: always colourable;
- $q \geq \Delta - \sqrt{\Delta} + 3$: polynomial-time (Molloy, Reed ’01 ’14);
- $q \leq \Delta - \sqrt{\Delta} + c$: NP-hard (Embden-Weinert, Hougardy, and Kreuter ’98).
Things can be more complicated!

On a planar graph, determining whether a graph is $q$-colourable:

- $q = 2$: easy;
- $q = 3$: NP-hard (Dailey ’80);
- $q = 4$: quadratic time (Four colour theorem) by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geq 5$: linear time (much simpler proof) (RSST ’96).
Things can be more complicated!

On a planar graph, determining whether a graph is $q$-colourable:

- $q = 2$: easy;
- $q = 3$: NP-hard (Dailey ’80);
- $q = 4$: quadratic time (Four colour theorem) by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geq 5$: linear time (much simpler proof) (RSST ’96).
Things can be more complicated!

On a planar graph, determining whether a graph is $q$-colourable:

- $q = 2$: easy;
- $q = 3$: NP-hard (Dailey ’80);
- $q = 4$: quadratic time (Four colour theorem) by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geq 5$: linear time (much simpler proof) (RSST ’96).
Properly colour a planar graph

Things can be more complicated!

On a planar graph, determining whether a graph is $q$-colourable:

- $q = 2$ : easy;
- $q = 3$ : $\text{NP}$-hard (Dailey ’80);
- $q = 4$ : quadratic time (Four colour theorem) by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geq 5$ : linear time (much simpler proof) (RSST ’96).
Things can be more complicated!

On a planar graph, determining whether a graph is $q$-colourable:

- $q = 2$: easy;
- $q = 3$: \textbf{NP}-hard (Dailey '80);
- $q = 4$: quadratic time (Four colour theorem) by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geq 5$: linear time (much simpler proof) (RSST '96).
What about generate a random proper colouring?

- $q > 2\Delta$ : rapid mixing of Glauber dynamics by Jerrum (1995);
- $q > \frac{11}{6}\Delta$ : rapid mixing of WSK dynamics by Vigoda (2000);
- $q < \Delta$ : NP-hard by Galanis, Štefankovič, and Vigoda (2015);
  (even $q$)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite $\Delta$-regular tree (namely a Bethe lattice), by Jonasson (2002).
What about generate a random proper colouring?

- $q > 2\Delta$: rapid mixing of Glauber dynamics by Jerrum (1995);
- $q > \frac{11}{6}\Delta$: rapid mixing of WSK dynamics by Vigoda (2000);
- $q < \Delta$: NP-hard by Galanis, Štefankovič, and Vigoda (2015); (even $q$)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite $\Delta$-regular tree (namely a Bethe lattice), by Jonasson (2002).
What about generate a random proper colouring?

- $q > 2\Delta$: rapid mixing of Glauber dynamics by Jerrum (1995);
- $q > \frac{11}{6}\Delta$: rapid mixing of WSK dynamics by Vigoda (2000);
- $q < \Delta$: NP-hard by Galanis, Štefankovič, and Vigoda (2015); (even $q$)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite $\Delta$-regular tree (namely a Bethe lattice), by Jonasson (2002).
What about generate a random proper colouring?

- \( q > 2\Delta \): rapid mixing of Glauber dynamics by Jerrum (1995);
- \( q > \frac{11}{6}\Delta \): rapid mixing of WSK dynamics by Vigoda (2000);
- \( q < \Delta \): NP-hard by Galanis, Štefankovič, and Vigoda (2015); (even \( q \))

It is conjectured that there is a threshold and \( q_c = \Delta + 1 \). This is the uniqueness threshold of Gibbs measures in an infinite \( \Delta \)-regular tree (namely a Bethe lattice), by Jonasson (2002).
What about generate a random proper colouring?

- $q > 2\Delta$: rapid mixing of Glauber dynamics by Jerrum (1995);
- $q > \frac{11}{6}\Delta$: rapid mixing of WSK dynamics by Vigoda (2000);
- $q < \Delta$: $\text{NP}$-hard by Galanis, Štefankovič, and Vigoda (2015);
  (even $q$)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite $\Delta$-regular tree (namely a Bethe lattice), by Jonasson (2002).
Sometimes you just cannot let it go.

\[ q = \Delta + 1 = 4 \]

credit: Chihao Zhang
Markov chain is a random walk in the solution space. (The solution space has to be connected!)
A disconnected state space is really bad.
There is still some hope if a giant component dominates.
There is still some hope if a giant component dominates. Approximately counting 4-colourings in cubic graphs by Lu, Yang, Zhang, Zhu (2017). (Not via Markov chains!)
A proper hypergraph colouring is one where no edge is monochromatic.
For \( k \)-uniform hypergraphs, Bordewich, Dyer, and Karpinski (2006) show that Glauber dynamics is rapidly mixing if

\[
\text{\( k \geq 4 \) and \( q > \Delta \) \quad \text{or} \quad \text{\( k = 3 \) and \( q > 1.5\Delta \).}
\]

However, Lovász local lemma implies that there exists a proper colouring if \( q > 2e\Delta^{1/(k-1)} \).

Frieze and Melsted (2011) showed that if \( q \ll \Delta \), then there exists a colouring so that no move is possible ("frozen").

Frieze and Anastasos (2017) showed that Glauber dynamics still converges rapidly if the hypergraph is simple and \( q > \max\{C_k \log n, 500k^3 \Delta^{1/(k-1)}\} \).

(Simple: every two hyperedges intersect with at most one vertex.)
Previous results

For k-uniform hypergraphs, Bordewich, Dyer, and Karpinski (2006) show that Glauber dynamics is rapidly mixing if

\[ k \geq 4 \text{ and } q > \Delta \quad \text{or} \quad k = 3 \text{ and } q > 1.5\Delta. \]

However, Lovász local lemma implies that there exists a proper colouring if \( q > 2e\Delta^{1/(k-1)} \).

Frieze and Melsted (2011) showed that if \( q \ll \Delta \), then there exists a colouring so that no move is possible (“frozen”).

Frieze and Anastos (2017) showed that Glauber dynamics still converges rapidly if the hypergraph is \text{simple} and \( q > \max\{C_k \log n, 500k^3 \Delta^{1/(k-1)}\} \).

\text{(Simple: every two hyperedges intersect with at most one vertex.)}
For $k$-uniform hypergraphs, Bordewich, Dyer, and Karpinski (2006) show that Glauber dynamics is rapidly mixing if

$$k \geq 4 \text{ and } q > \Delta \text{ or } k = 3 \text{ and } q > 1.5\Delta.$$  

However, Lovász local lemma implies that there exists a proper colouring if $q > 2e\Delta^{1/(k-1)}$.

Frieze and Melsted (2011) showed that if $q \ll \Delta$, then there exists a colouring so that no move is possible (“frozen”).

Frieze and Anastos (2017) showed that Glauber dynamics still converges rapidly if the hypergraph is simple and $q > \max\{C_k \log n, 500k^3\Delta^{1/(k-1)}\}$. (Simple: every two hyperedges intersect with at most one vertex.)
For k-uniform hypergraphs, Bordewich, Dyer, and Karpinski (2006) show that Glauber dynamics is rapidly mixing if

\[ k \geq 4 \text{ and } q > \Delta \quad \text{or} \quad k = 3 \text{ and } q > 1.5\Delta. \]

However, Lovász local lemma implies that there exists a proper colouring if \( q > 2e\Delta^{1/(k-1)} \).

Frieze and Melsted (2011) showed that if \( q \ll \Delta \), then there exists a colouring so that no move is possible ("frozen").

Frieze and Anastasos (2017) showed that Glauber dynamics still converges rapidly if the hypergraph is simple and \( q > \max\{C_k \log n, 500k^3\Delta^{1/(k-1)}\} \).

(Simple: every two hyperedges intersect with at most one vertex.)
Our results

Theorem

For integers $\Delta \geq 2$, $k \geq 28$, and $q > 315\Delta^{\frac{14}{k-14}}$, there is an FPTAS for $q$-colourings in $k$-uniform hypergraphs with maximum degree $\Delta$.

Theorem

For integers $\Delta \geq 2$, $k \geq 28$, and $q > 798\Delta^{\frac{16}{k-16/3}}$, there is also an almost-uniform polynomial-time sampler.

Our approach is based on a result by Moitra (2017), whose original approach in this setting would require $k > C \log \Delta$ in addition.
Theorem

For integers \( \Delta \geq 2 \), \( k \geq 28 \), and \( q > 315\Delta^{\frac{14}{k-14}} \), there is an FPTAS for \( q \)-colourings in \( k \)-uniform hypergraphs with maximum degree \( \Delta \).

Theorem

For integers \( \Delta \geq 2 \), \( k \geq 28 \), and \( q > 798\Delta^{\frac{16}{k-16/3}} \), there is also an almost-uniform polynomial-time sampler.

Our approach is based on a result by Moitra (2017), whose original approach in this setting would require \( k > C \log \Delta \) in addition.
Theorem

For integers $\Delta \geq 2$, $k \geq 28$, and $q > 315\Delta^{\frac{14}{k-14}}$, there is an FPTAS for $q$-colourings in $k$-uniform hypergraphs with maximum degree $\Delta$.

Theorem

For integers $\Delta \geq 2$, $k \geq 28$, and $q > 798\Delta^{\frac{16}{k^{16/3}}}$, there is also an almost-uniform polynomial-time sampler.

Our approach is based on a result by Moitra (2017), whose original approach in this setting would require $k > C \log \Delta$ in addition.
Tensor perspective

Each vertex is a **Equality** tensor, $q$-dimensional, order $\leq \Delta$.

Each hyperedge is a **Colouring** tensor, $q$-dimensional, order $k$.

Thus, the hypergraph can be viewed as a **bipartite** tensor network. The total number of proper colourings is the contraction of the whole tensor network.
Each vertex is an **EQUALITY** tensor, $q$-dimensional, order $\leq \Delta$.

Each hyperedge is a **COLOURING** tensor, $q$-dimensional, order $k$.

Thus, the hypergraph can be viewed as a **bipartite** tensor network. The total number of proper colourings is the contraction of the whole tensor network.
MOITRA’S APPROACH

(WITH OUR MODIFICATIONS)
Lovász Local Lemma

Let $H = (V, E)$ be the hypergraph, and $\Gamma(e)$ be neighbourhood of $e \in E$.

**Theorem (Lovász '77)**

If there exists an assignment $x : E \to (0, 1)$ such that for every $e \in E$ we have

$$\Pr(e \text{ is monochromatic}) \leq x(e) \prod_{e' \in \Gamma(e)} (1 - x(e')),$$

then a proper colouring exists.

Normally we set $x(e) = \frac{1}{k\Delta}$ and the condition becomes $q > (ek\Delta)^{1 \over k-1}$.

Let $\mu_e(\cdot)$ be the Gibbs (uniform) distribution on all proper colourings.

**Theorem (Haeupler, Saha, and Srinivasan '11)**

If (1) holds for every $e \in E$, then for any event $B$, it holds that

$$\mu_e(B) \leq \Pr(B) \prod_{e \in \Gamma(B)} (1 - x(e))^{-1}.$$
Let $H = (V, E)$ be the hypergraph, and $\Gamma(e)$ be neighbourhood of $e \in E$.

**Theorem (Lovász ’77)**

If there exists an assignment $x : E \to (0, 1)$ such that for every $e \in E$ we have

$$\Pr(e \text{ is monochromatic}) \leq x(e) \prod_{e' \in \Gamma(e)} (1 - x(e')),$$

then a proper colouring exists.

Normally we set $x(e) = \frac{1}{k\Delta}$ and the condition becomes $q > (e k\Delta) \frac{1}{k-1}$.

Let $\mu_e(\cdot)$ be the Gibbs (uniform) distribution on all proper colourings.

**Theorem (Haeupler, Saha, and Srinivasan ’11)**

If (1) holds for every $e \in E$, then for any event $B$, it holds that

$$\mu_e(B) \leq \Pr(B) \prod_{e \in \Gamma(B)} (1 - x(e))^{-1}.$$
Let $H = (V, \mathcal{E})$ be the hypergraph, and $\Gamma(e)$ be neighbourhood of $e \in \mathcal{E}$.

**Theorem (Lovász ’77)**

*If there exists an assignment $\chi : \mathcal{E} \to (0, 1)$ such that for every $e \in \mathcal{E}$ we have*

$$
\Pr(e \text{ is monochromatic}) \leq \chi(e) \prod_{e' \in \Gamma(e)} (1 - \chi(e')),
$$

*then a proper colouring exists.*

Normally we set $\chi(e) = \frac{1}{k\Delta}$ and the condition becomes $q > (ek\Delta)^{\frac{1}{k-1}}$.

Let $\mu_e(\cdot)$ be the Gibbs (uniform) distribution on all proper colourings.

**Theorem (Haeupler, Saha, and Srinivasan ’11)**

*If (1) holds for every $e \in \mathcal{E}$, then for any event $B$, it holds that*

$$
\mu_e(B) \leq \Pr(B) \prod_{e \in \Gamma(B)} (1 - \chi(e))^{-1}.
$$
Lemma

If $t \geq k$ and $q \geq (et\Delta)^{\frac{k-1}{k}}$, then for any $v \in V$ and any colour $c \in [q],$

$$\frac{1}{q} \left(1 - \frac{1}{t}\right) \leq \Pr_{\sigma \sim \mu_c} (\sigma(v) = c) \leq \frac{1}{q} \left(1 + \frac{4}{t}\right).$$

We use this lemma with $t \approx \Delta^C$ at various places with various $C$. Recall that our assumption is of the form $q \geq C'\Delta^C\Delta^{c''}$. Under $\mu_c$, all vertices are very close to uniform.

The use of this lemma is when some vertices are already coloured, namely $\mu_c$ conditioned on a partial colouring. Replace $k$ with the minimum number of uncoloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is $\frac{\varepsilon}{n}$-approximation of the marginals.
Lemma

If \( t \geq k \) and \( q \geq (et\Delta)^{\frac{1}{k-1}} \), then for any \( v \in V \) and any colour \( c \in [q] \),

\[
\frac{1}{q} \left( 1 - \frac{1}{t} \right) \leq \Pr_{\sigma \sim \mu_c} (\sigma(v) = c) \leq \frac{1}{q} \left( 1 + \frac{4}{t} \right).
\]

We use this lemma with \( t \approx \Delta^C \) at various places with various \( C \). Recall that our assumption is of the form \( q \geq C' \Delta^C \frac{C''}{k} \).

Under \( \mu_c \), all vertices are very close to uniform.

The use of this lemma is when some vertices are already coloured, namely \( \mu_c \) conditioned on a partial colouring. Replace \( k \) with the minimum number of uncoloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is \( \frac{\epsilon}{n} \)-approximation of the marginals.
**Locally uniformity**

**Lemma**

If $t \geq k$ and $q \geq (e t \Delta)^{\frac{1}{k-1}}$, then for any $v \in V$ and any colour $c \in [q]$,

$$\frac{1}{q} \left(1 - \frac{1}{t}\right) \leq P_{\sigma \sim \mu_c} (\sigma(v) = c) \leq \frac{1}{q} \left(1 + \frac{4}{t}\right).$$

We use this lemma with $t \approx \Delta^C$ at various places with various $C$. Recall that our assumption is of the form $q \geq C' \Delta^\frac{C''}{k}$.

Under $\mu_c$, all vertices are very close to uniform.

The use of this lemma is when some vertices are already coloured, namely $\mu_c$ conditioned on a partial colouring. Replace $k$ with the minimum number of uncoloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is $\frac{\epsilon}{n}$-approximation of the marginals.
Lemma

If $t \geq k$ and $q \geq (et\Delta)^{\frac{1}{k-1}}$, then for any $v \in V$ and any colour $c \in [q]$,

$$\frac{1}{q} \left(1 - \frac{1}{t}\right) \leq \Pr_{\sigma \sim \mu_c} (\sigma(v) = c) \leq \frac{1}{q} \left(1 + \frac{4}{t}\right).$$

We use this lemma with $t \approx \Delta^c$ at various places with various $C$. Recall that our assumption is of the form $q \geq C'\Delta^{c''}/k$.

Under $\mu_c$, all vertices are very close to uniform.

The use of this lemma is when some vertices are already coloured, namely $\mu_c$ conditioned on a partial colouring. Replace $k$ with the minimum number of un-coloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is $\frac{\epsilon}{n}$-approximation of the marginals.
**Locally uniformity**

**Lemma**

If $t \geq k$ and $q \geq (e t \Delta)^{\frac{1}{k-1}}$, then for any $v \in V$ and any colour $c \in [q]$,

$$\frac{1}{q} \left(1 - \frac{1}{t}\right) \leq \Pr_{\sigma \sim \mu_c} (\sigma(v) = c) \leq \frac{1}{q} \left(1 + \frac{4}{t}\right).$$

We use this lemma with $t \approx \Delta^C$ at various places with various $C$. Recall that our assumption is of the form $q \geq C' \Delta^C \frac{C''}{k}$.

Under $\mu_c$, all vertices are very close to uniform.

The use of this lemma is when some vertices are already coloured, namely $\mu_c$ conditioned on a partial colouring. Replace $k$ with the minimum number of uncoloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is $\frac{\varepsilon}{n}$-approximation of the marginals.
Let $C_i$ be the set of colourings where $v$ is coloured $c_i$, and $\mu_i$ be uniform over $C_i$. We want to couple $C_1$ and $C_2$.

Start: $V_1 = \{v\}$, $V_{col} = \{v\}$. Maintain $V_2 = V \setminus V_1$.

Body:

1. For any hyperedge $e$ intersecting both $V_1$ and $V_2$, let $u$ be its first vertex. Couple $u$ maximally assuming we know its marginal probabilities.

2. Remove all hyperedges that are satisfied in both copies.

3. If an edge has $k_1$ vertices coloured, put other vertices in $V_1$ (failed) and remove it.

Stop: all hyperedges intersecting both $V_1$ and $V_2 \setminus V_{col}$ are satisfied.

If $q > C \Delta^{\frac{3}{k-k_1}}$, then the coupling stops in $O(\log n)$ steps with probability $1 - O(\frac{1}{n^c})$. 
Let $C_i$ be the set of colourings where $v$ is coloured $c_i$, and $\mu_i$ be uniform over $C_i$. We want to couple $C_1$ and $C_2$.

**Start:** $V_1 = \{v\}$, $V_{col} = \{v\}$. Maintain $V_2 = V \setminus V_1$.

**Body:**

1. For any hyperedge $e$ intersecting both $V_1$ and $V_2$, let $u$ be its first vertex. Couple $u$ maximally assuming we know its marginal probabilities.

2. Remove all hyperedges that are satisfied in both copies.

3. If an edge has $k_1$ vertices coloured, put other vertices in $V_1$ (failed) and remove it.

**Stop:** all hyperedges intersecting both $V_1$ and $V_2 \setminus V_{col}$ are satisfied.

If $q > \mathcal{C} \Delta \frac{3}{k-k_1}$, then the coupling stops in $O(\log n)$ steps with probability $1 - O\left(\frac{1}{n^c}\right)$. 
Let $C_i$ be the set of colourings where $v$ is coloured $c_i$, and $\mu_i$ be uniform over $C_i$. We want to couple $C_1$ and $C_2$.

Start: $V_1 = \{v\}$, $V_{col} = \{v\}$. Maintain $V_2 = V \setminus V_1$.

Body:

1. For any hyperedge $e$ intersecting both $V_1$ and $V_2$, let $u$ be its first vertex. Couple $u$ maximally assuming we know its marginal probabilities.
2. Remove all hyperedges that are satisfied in both copies.
3. If an edge has $k_1$ vertices coloured, put other vertices in $V_1$ (failed) and remove it.

Stop: all hyperedges intersecting both $V_1$ and $V_2 \setminus V_{col}$ are satisfied.

If $q > C \Delta^{k-1} k^{3}$, then the coupling stops in $O(\log n)$ steps with probability $1 - O\left(\frac{1}{n^c}\right)$. 

Coupling
Coupling tree $\mathcal{T}$: each node is a pair of partial colourings $(x, y)$.

The children of $(x, y)$ are all $q^2$ ways to extend them to the next vertex. All information can be recovered by simulating the coupling from the start.

Moitra (2017) marks what vertices to couple in advance, whereas our coupling is adaptive.
Coupling tree $\mathcal{T}$: each node is a pair of partial colourings $(x, y)$.

The children of $(x, y)$ are all $q^2$ ways to extend them to the next vertex. All information can be recovered by simulating the coupling from the start.

Moitra (2017) marks what vertices to couple in advance, whereas our coupling is adaptive.
Coupling tree $\mathcal{T}$: each node is a pair of partial colourings $(x, y)$.

The children of $(x, y)$ are all $q^2$ ways to extend them to the next vertex. All information can be recovered by simulating the coupling from the start.

Moitra (2017) marks what vertices to couple in advance, whereas our coupling is adaptive.
Let \( q_i = \frac{|c_i|}{|C|} \) for \( i = 1, 2 \).

A sampler \( S \):

1. Sample \((X, Y)\) using the coupling;
2. Uniformly output a colouring in \( C_X \) with probability \( \frac{q_1}{q_1 + q_2} \). Otherwise uniformly output a colouring in \( C_Y \).

The output of \( S \) is uniform over \( C_1 \cup C_2 \).
We cannot really run the coupling. Instead, we set up a linear program. The variables are to mimic:

\[ p^x_{x,y} = q_1 \cdot \frac{|C_1 \cup C_2|}{|C_x|} \cdot \mu_{cp}(x, y); \]
\[ p^y_{x,y} = q_2 \cdot \frac{|C_1 \cup C_2|}{|C_y|} \cdot \mu_{cp}(x, y); \]

\[ 0 \leq p^x_{x,y}, p^y_{x,y} \leq 1. \]

The meaning of \( p^x_{x,y} \) is, for any \( \sigma \in C_x \),

“conditioned on \( \mathcal{S} \) outputing \( \sigma \), the coupling reaches \( (x, y) \)”.

This definition is independent from \( \sigma \).
We cannot really run the coupling. Instead, we set up a linear program. The variables are to mimic:

\[
p^x_{x,y} = q_1 \cdot \frac{|C_1 \cup C_2|}{|C_x|} \cdot \mu_{cp}(x, y);
\]

\[
p^y_{x,y} = q_2 \cdot \frac{|C_1 \cup C_2|}{|C_y|} \cdot \mu_{cp}(x, y);
\]

\[0 \leq p^x_{x,y}, p^y_{x,y} \leq 1.\]

The meaning of \( p^x_{x,y} \) is, for any \( \sigma \in C_x \),

“conditioned on \( S \) outputing \( \sigma \), the coupling reaches \((x, y)\)”.

This definition is independent from \( \sigma \).
From the definition: \( \frac{p_{x,y}^x}{p_{x,y}^y} = \frac{q_1}{q_2} \cdot \frac{|c_y|}{|c_x|} \).

If \((x, y)\) is a leaf in \(T\), then we can compute \( \frac{|c_x|}{|c_y|} \) in time \( \exp(|V_1|) \).

**Constraints 1:** For every leaf \((x, y)\), we have the constraints:

\[
\frac{p_{x,y}^x}{p_{x,y}^y} \cdot \frac{|c_x|}{|c_y|} \leq r.
\]

Here \(r\) and \(\bar{r}\) are our guessed lower and upper bounds for \(\frac{q_1}{q_2}\).
From the definition: \[
\frac{p^x_{x,y}}{p^y_{x,y}} = \frac{q_1}{q_2} \cdot \frac{|c_y|}{|c_x|}.
\]

If \((x, y)\) is a leaf in \(T\), then we can compute \(\frac{|c_x|}{|c_y|}\) in time \(\exp(|V_1|)\).

**Constraints 1**: For every leaf \((x, y)\), we have the constraints:

\[
r \leq \frac{p^x_{x,y}}{p^y_{x,y}} \cdot \frac{|c_x|}{|c_y|} \leq \bar{r}.
\]

Here \(r\) and \(\bar{r}\) are our guessed lower and upper bounds for \(\frac{q_1}{q_2}\).
From the definition: \( \frac{p_{x,y}^x}{p_{x,y}^y} = \frac{q_1}{q_2} \cdot \frac{|c_y|}{|c_x|} \).

If \((x, y)\) is a leaf in \(\mathcal{T}\), then we can compute \(\frac{|c_x|}{|c_y|}\) in time \(\exp(|V_1|)\).

**Constraints 1**: For every leaf \((x, y)\), we have the constraints:

\[
\bar{r} \leq \frac{p_{x,y}^x}{p_{x,y}^y} \cdot \frac{|c_x|}{|c_y|} \leq \bar{r}.
\]

Here \(\bar{r}\) and \(\bar{r}\) are our guessed lower and upper bounds for \(\frac{q_1}{q_2}\).
Constraints 2: For the root $(x_0, y_0) \in \mathcal{I}$, we have

$$p_{x_0, y_0} = p_{x_0, y_0} = 1.$$ 

Moreover, for every non-leaf $(x, y) \in \mathcal{I}$, let $u$ be the next vertex to couple. We have:

for every $c \in [q]$,  

$$p_{x, y} = \sum_{c' \in [q]} p_{x_{u \leftarrow c}, y_{u \leftarrow c'));$$

for every $c \in [q]$,  

$$p_{x, y} = \sum_{c' \in [q]} p_{x_{u \leftarrow c'}, y_{u \leftarrow c}.$$
Due to **Constraints 2**, a simple induction shows that for every $\sigma \in C_1$,

$$
\sum_{(x,y) \in \mathcal{L}(T): \sigma \models x} p_{x,y}^x = 1.
$$

Rewrite $|C_1|$:

$$
|C_1| = \sum_{\sigma \in C_1} 1 = \sum_{\sigma \in C_1} \sum_{(x,y) \in \mathcal{L}(T): \sigma \models x} p_{x,y}^x
= \sum_{(x,y) \in \mathcal{L}(T)} \sum_{\sigma \models x} p_{x,y}^x
= \sum_{(x,y) \in \mathcal{L}(T)} p_{x,y}^x |C_x|.
$$

Similar equalities hold on the $Y$ side, implying:

$$
\frac{q_1}{q_2} = \frac{|C_1|}{|C_2|} = \frac{\sum_{(x,y) \in \mathcal{L}(T)} p_{x,y}^x |C_x|}{\sum_{(x,y) \in \mathcal{L}(T)} p_{x,y}^y |C_y|}.
$$
Recover the marginals (cont.)

\[
\frac{q_1}{q_2} = \frac{\sum_{(x, y) \in \mathcal{L}(\mathcal{J})} p_{x,y}^x |C_x|}{\sum_{(x, y) \in \mathcal{L}(\mathcal{J})} p_{x,y}^y |C_y|}
\]

Recall **Constraints 1.** For any \((x, y) \in \mathcal{L}(\mathcal{J}),\)

\[
\frac{p_{x,y}^x |C_x|}{p_{x,y}^y |C_y|} \leq r \leq \bar{r}.
\]

It implies that

\[
\frac{q_1}{q_2} \leq \frac{q_1}{q_2} \leq \bar{r}.
\]
Unfortunately, the whole linear program is exponentially large, but the coupling stops at $O(\log n)$ size whp. If we truncate at $O(\log n)$ levels, the error should be small, due to local uniformity.

**Constraints 3:** For every $c, c' \in [q]$ that $c \neq c'$:

$$p_{x \leftarrow c, y \leftarrow c'}^{x \leftarrow c} \leq \frac{5}{t} \cdot p_{x,y}^{x};$$

$$p_{x \leftarrow c, y \leftarrow c'}^{y \leftarrow c'} \leq \frac{5}{t} \cdot p_{x,y}^{y}.$$

The quantity $t$ will eventually be set as $C(k\Delta)^6$. 
Unfortunately, the whole linear program is exponentially large, but the coupling stops at $O(\log n)$ size whp. If we truncate at $O(\log n)$ levels, the error should be small, due to local uniformity.

**Constraints 3:** For every $c, c' \in [q]$ that $c \neq c'$:

\[
p^{x \leftarrow c}_{x u \leftarrow c, y u \leftarrow c'} \leq \frac{5}{t} \cdot p^x_{x, y}; \\
p^{y \leftarrow c'}_{x u \leftarrow c, y u \leftarrow c'} \leq \frac{5}{t} \cdot p^y_{x, y}.
\]

The quantity $t$ will eventually be set as $C(k\Delta)^6$. 
Recall that

$$|\mathcal{C}_1| = \sum_{\sigma \in \mathcal{C}_1} \sum_{(x, y) \in \mathcal{L}(\mathcal{I}) : \sigma \models x} p_{x, y}^x. $$

The truncation error from a particular $\sigma \in \mathcal{C}_1$ comes from conditioned on outputing $\sigma$, the coupling lasts too long.

Such “bad” colouring does exist (all early vertices are monochromatic).

We prove two things:

1. The fraction of “bad” colourings is small;
2. For every “good” colouring, the truncation error is small because of Constraints 3.
Recall that

\[ |\mathcal{C}_1| = \sum_{\sigma \in \mathcal{C}_1} \sum_{(x,y) \in \mathcal{L}(I): |x|=x} p_{x,y}^x. \]

The truncation error from a particular \( \sigma \in \mathcal{C}_1 \) comes from conditioned on outputing \( \sigma \), the coupling lasts too long.

Such “bad” colouring does exist (all early vertices are monochromatic).

We prove two things:

1. The fraction of “bad” colourings is small;
2. For every “good” colouring, the truncation error is small because of Constraints 3.
Recall that

\[ |\mathcal{C}_1| = \sum_{\sigma \in \mathcal{C}_1} \sum_{(x,y) \in \mathcal{L}(I): \sigma|_x = x} p_{x,y}^x. \]

The truncation error from a particular \( \sigma \in \mathcal{C}_1 \) comes from conditioned on outputing \( \sigma \), the coupling lasts too long.

Such “bad” colouring does exist (all early vertices are monochromatic).

We prove two things:

1. The fraction of “bad” colourings is small;
2. For every “good” colouring, the truncation error is small because of **Constraints 3**.
So far we are calculating the marginal probability, which requires that there are sufficiently many vertices in all hyperedges.

- For approximate counting, we use the local lemma again, to find an assignment so that every hyperedge is satisfied by the first \( \frac{k}{14} \) vertices. Then we compute the marginal probability of this assignment by fixing vertices one by one.

- For sampling, we use the marginal to colour vertices, similar to the coupling process. With high probability, the remaining connected components are of size \( O(\log n) \).
So far we are calculating the marginal probability, which requires that there are sufficiently many vertices in all hyperedges.

• For approximate counting, we use the local lemma again, to find an assignment so that every hyperedge is satisfied by the first \( \frac{k}{14} \) vertices. Then we compute the marginal probability of this assignment by fixing vertices one by one.

• For sampling, we use the marginal to colour vertices, similar to the coupling process. With high probability, the remaining connected components are of size \( O(\log n) \).
So far we are calculating the marginal probability, which requires that there are sufficiently many vertices in all hyperedges.

- For approximate counting, we use the local lemma again, to find an assignment so that every hyperedge is satisfied by the first $\frac{k}{14}$ vertices. Then we compute the marginal probability of this assignment by fixing vertices one by one.

- For sampling, we use the marginal to colour vertices, similar to the coupling process. With high probability, the remaining connected components are of size $O(\log n)$. 
CONCLUDING REMARKS
Open Problems

- What is the correct threshold for hypergraph colouring?
  - Is it \( q \approx \Delta^{\frac{2}{k}} \)?

- What about \textbf{NP}-hardness of sampling hypergraph colourings?

- Does this method work for general LLL?

- What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?
Open Problems

• What is the correct threshold for hypergraph colouring?
  • Is it $q \approx \Delta^{2/k}$?

• What about \textbf{NP}-hardness of sampling hypergraph colourings?

• Does this method work for general LLL?

• What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?
Open Problems

• What is the correct threshold for hypergraph colouring?
  • Is it $q \approx \Delta^{2/k}$?

• What about \textbf{NP}-hardness of sampling hypergraph colourings?

• Does this method work for general LLL?

• What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?


Open problems

- What is the correct threshold for hypergraph colouring?
  - Is it $q \approx \Delta^{\frac{2}{k}}$?

- What about \textbf{NP}-hardness of sampling hypergraph colourings?

- Does this method work for general LLL?

- What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?
Open problems

• What is the correct threshold for hypergraph colouring?
  • Is it $q \approx \Delta \frac{2}{k}$?

• What about NP-hardness of sampling hypergraph colourings?

• Does this method work for general LLL?

• What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?
A professor is one who can speak on any subject for precisely fifty minutes.

— Norbert Wiener

THANK YOU!
arXiv:1711.03396